DOCUMENT RESUME

ED 047 960 SE 010 799

AUTHOR	Gay, Lorraine R.
TITLE	Temporal Position of Reviews and Its Effect on the Retention of Mathematical Rules.
INSTITUTION	Florida State Univ., Tallahassee.
PUB DATE	Nov 70
NOTE	85p.; Paper presented at the Annual Meeting of the American Educational Research Association (Feb. 4-7, 1971, New York, N.Y.)
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	*Achievement, Algebra, Doctoral Theses, *Elementary School Mathematics, Geometry, *Instruction, *Learning, Retention, *Time Factors (Learning)

ABSTRACT

Reported are the results of a study which investigated the effect of temporal position of review through two experiments. Experiment I involved one review; Experiment II involved two reviews. In Experiment I, eighth grade students were pretested, presented with the learning task (four rules from algebra and geometry), and randomly assigned to one of four groups. Reviews were presented one day (Group I), one week (II), and two weeks (III) after criterion. The control group got no review. Three weeks after the original learning, a retention test was administered. Although the review groups retained more than the control group, temporal position was not significant. Seventh grade students were used in Experiment II. Reviews were made after one and two days (Group I), one and seven days (II), or six and seven days (III). The control group received no review. A three-week retention test showed that the review groups retained significantly more than the control group. Group II retained significantly more than Group I. It was concluded that one review is more effective than no review, while an early and a late review produce optimal retention over a three-week interval. (Author/RS)

ED047960

TEMPORAL POSITION OF REVIEWS AND ITS EFFECT ON THE RETENTION OF MATHEMATICAL RULES

By

LORRAINE R. GAY

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

FLORIDA STATE UNIVERSITY
Tallahassee, Florida
November, 1970

SE 010 799

ABSTRACT

The purpose of the present study was to investigate for mathematical rule learning one variable, spaced review, which verbal learning studies have indicated is important for retention. The effect of temporal position of review was investigated in two experiments: experiment I involved one review; experiment II involved two reviews. The methodology for both experiments paralleled that employed in meaningful verbal learning research.

Experiment I compared the relative effectiveness of one review placed either one day (Group I), one week (Group II), or two weeks (Group III) after original learning. Learning consisted of four rules presented via computer-assisted instruction. Subjects utilized varying numbers of examples, all learning to a common criterion of two consecutive correct problem solutions. The four rules involved principles from algebra and geometry; raising an algebraic expression to an indicated power, determining the measure of the third angle of a triangle, determining the exponent of the product of factors and finding the geometric mean.

Fifty-three eighth grade students were pretested, presented with the learning task and randomly assigned to one of four groups; the four groups were the three groups described above plus a control group which received no review. The review procedure was very similar to the learning procedure differing only in that it began with the presentation of two

new problems. All subjects relearned to a criterion of two successful problem solutions for each rule.

Three weeks from the day of original learning, all subjects were administered a delayed retention measure on the four rules. While all review groups retained considerably more than the no review group, temporal position was not significant. These findings are consistent with previous research on meaningful verbal learning.

Experiment II compared the relative effectiveness of two reviews placed either one and two days (Group I), one and seven days (Group II) or six and seven days (Group III) after original learning. Sixty-seven seventh grade students were pretested, presented with the learning task, and randomly assigned to one of four groups; the four groups were the three groups just described plus a control group which received no review. The learning and review procedures were identical to those described for experiment I, all subjects learning to a common criterion.

Three weeks from the day of learning all subjects were administered a delayed retention measure on the four rules, the same one used in experiment I. All review groups retained significantly more than the no review group. Also the group with both an early and a delayed review retained significantly more than the group with two early reviews. This suggests that while an early and a late review each make a unique contribution to retention, the contribution of the latter is greater.

It was also found that in general, subjects in all groups required the same number of examples to reach criterion at the time of review 1 as at the time of original learning. At the time of review 1, however,

example latency was approximately 50 percent shorter than at the time of learning. In contrast, at the time of review 2, subjects required half as many examples to reach criterion as at the time of learning, with example latency being 75 percent shorter.

It was concluded that with respect to the retention of mathematical rules, one review is more effective than no review, regardless of temporal position. Optimal retention over a three week interval, however, is obtained with two reviews, one early and one delayed.

TABLE OF CONTENTS

		Page
ABSTRACT		iii
LIST OF TABLES		viii
LIST OF ILLUSTRATIONS		ix
CHAPTER		
I.	INTRODUCTION	1
	Significance of the Problem	
	Survey of the Literature on Review	
	Non-Meaningfully Related Material	
	Meaningfully Related Material	
	Effectiveness of Spaced Review	
	Temporal Position of Spaced Review	
II.	DESCRIPTION OF THE PRESENT STUDY	8
	Purpose and Rationale of the Study	
	Method and Procedure	
	Learning Materials	
	Apparatus	
	Selection of Subjects	
	Experimental Design	
	Experimental Procedures	
	Experiment I	
	Experiment II	
III.	ANALYSIS OF THE DATA	19
	Scoring Procedures	
	Retention Data	
	Experiment I	
	Experiment II	
	Item Difficulty	
	Rule Mastery	
	Group by Rule Repeated Measures Analysis	
	Within-Program Variables	
	Number of Examples	
	Length of Example Latency	
	Rule Difficulty and Example Latency	
	Correct Responses	
IV.	DISCUSSION	43
V.	SUMMARY	49

TABLE OF CONTENTS - continued

APPENDIX
A. LEARN PRETEST 53
B. POST-LEARN RETENTION TEST 55
C. PRE-REVIEW 1 RETENTION TEST 57
D. POST-REVIEW 1 RETENTION TEST 59
E. PRE-REVIEW 2 RETENTION TEST 61
F. POST-REVIEW 2 RETENTION TEST 63
G. DELAYED RETENTION TEST 65

REFERENCES 67

ADDITIONAL REFERENCES 72

VITA 74

LIST OF TABLES

TABLE		Page
1.	Means and standard deviations for all tests by group and experiment	22
2.	Analysis of variance of delayed retention scores for review and no-review subjects for experiment II .	25
3.	Summary of Scheffé confidence interval procedure for a posteriori analysis of group means for experiment II	26
4.	Percent of students successfully solving each problem by group, session and experiment	28
5.	Percent of subjects achieving varying degrees of mastery on each of the four rules for experiment I and experiment II	29
6.	Percent of subjects achieving mastery at each session and the cumulative percent over successive session of retention measurement, for the four rules (R1, R2, R3, R4), for experiment I and experiment II .	31
7.	A group by rule repeated measures analysis for experiment I and experiment II	32
8.	A comparison of the number of examples viewed at each session for experiment I and experiment II . . .	33
9.	A comparison of mean example latency and rule difficulty for experiment I and experiment II	35
10.	A comparison of mean correct response latency and rule difficulty for experiment I and experiment II . .	41

LIST OF ILLUSTRATIONS

FIGURE		Page
1.	Experimental designs for experiment I and experiment II	14
2.	Plot of test scores for all groups before and after each learning and review session	23
3.	A comparison of three measures of retention for experiment I and experiment II	38

I. INTRODUCTION

One important area of application of learning research to instructional effectiveness is that of maximizing the retention of learned material. It has repeatedly been found that in many areas of learning there is considerable forgetting over periods of time (Schrepel & Laslett, 1936; Sterrett & Davis, 1954). Pressey, Robinson and Horrocks (1959) determined that approximately sixty-six percent of the concepts "learned" in high school and college courses are forgotten within two years. The problem is even more critical in subjects such as science (Tyler, 1930; Freud & Cheronis, 1940; Smeltz, 1956) and mathematics (Worchester, 1928; Layton, 1932; Lahey, 1941; Davis & Rood, 1947), curriculum areas which involve the learning of intellectual skills. Unlike verbal information, a forgotten skill cannot be regained through simple rereading. It is less than consoling that intellectual skills appear to be retained somewhat better than verbal information (Tyler, 1930; Freud & Cheronis, 1940; Davis & Rood, 1947), since neither may be retained to a satisfactory degree. For example, algebra is composed mainly of rules or algorithms, and yet as Layton (1932) has reported, up to 66% of first year algebra material may be non-retrievable after a period of one year.

Ausubel (1968) suggests that the retention losses described in many studies are not inevitable. A major portion of the loss, he states, is reflective of rote learning, of poorly organized and programmed subject matter, of correctable ambiguity and confusion in the presentation of ideas, and of inadequate pacing and review of material. He further states that if adequate attention were paid to such considerations as optimal review, students might retain over a lifetime most of the important ideas they learned in school.

Significance of the Problem

Review is a prevalent phenomenon in mathematics, as in all curriculum areas. Systematic investigation of review can provide valuable information to educators at all levels, from curriculum designers to classroom teachers. It would be of considerable worth to a teacher to know if review is worth the additional class time. The more important issue, however, is that of temporal position. It makes little difference to the teacher when the review takes place and it would be a simple matter to schedule review at the time(s) most conducive to improving retention. There are few factors in education where such a small initial difference in instruction can account for such a significant difference in future retention.

Survey of the Literature on Review

In order to formulate the problem to be investigated, it will be useful to review briefly, by reference to published reviews as well as representative studies, what the major findings have been regarding

reviews in the learning of non-meaningfully related materials. These findings are a kind of basic knowledge which may or may not be directly applicable to the problem of rule retention, but which can at least provide a contrasting background. Further, it will be of interest to examine the more relevant studies concerning the retention of meaningfully related material, mainly prose. These studies include those which address themselves to the question of the effectiveness of spaced review as well as those investigating the problem of temporal positioning of spaced review.

Until recently, the research emphasis has been on retention of non-meaningfully related materials over relatively short periods of time. The findings of such research are not always applicable to school learning: Ausubel (1957) provides some evidence which indicates that the phenomenon generally referred to as "retroactive inhibition" does not necessarily occur in meaningful learning. However, a number of studies (Dowling & Braun, 1957) have convincingly demonstrated that decline in retention and final level of retention are dependent on such factors as meaningfulness of the material and interrelatedness of its components, which in turn are a function of such factors as nature of the material, the degree of original learning, and review. Recent research (Schuell & Keppel, 1970) also indicates that while individual student abilities result in sizeable differences in amount of material retained, these differences are not apparent when equivalent levels of learning are involved. Mouly (1968) has stated that if students, regardless of learning rate, achieve the same degree of learning before the

retention interval is introduced, there is little evidence that their forgetting rates will differ beyond a negligible degree. What determines the level of forgetting is mainly the degree of original learning, regardless of the time and effort required to get the learning up to a given level of acquisition.

Non-Meaningfully Related Material

Studies of the learning of nonsense syllables and word pairs have found a strong indication of relationship between amount of repetition, or review, and learning (Melton, 1963; Tulving, 1968; Underwood, 1964). The majority of these studies have been concerned with retention over short periods of time. Such studies (Pratt, 1936; Murphy, 1953; Peterson, 1956; Korn & Lindley, 1963; Boschke & Lim, 1967) typically report frequency of repetition to be directly related to degree of recall and, to a lesser extent, to degree of recognition. Similarly, studies investigating non-meaningful material using both short-term and long-term retention paradigms (Kintz, 1965) have found amount of forgetting to be a function of repetition. Of course, excessive overlearning tends to be uneconomical (Shay, 1961); beyond a certain point further training efforts result in relatively small retention gains. Generally, studies in this field lead one to expect better retention as a result of increased repetition or practice, even when such practice is given immediately following the attainment of some criterion of initial learning.

Meaningfully Related Material

For meaningfully related material such as prose passages, it has often been found that practice in the form of spaced reviews rather than contiguous repetitions is particularly effective in promoting retention. This suggests, basically, that distributed practice is better than massed practice for insuring retention of meaningful material. Massed and distributed practice have two somewhat different meanings depending on the nature of the task. When the task is one of learning verbal pairs or longer verbal sequences, practice, whether massed or distributed, refers to the repeated presentation of the same material. When the task is one of concept or rule learning, however, amount of practice refers to number of different examples, whether these are presented contiguously as massed practice or as reviews spaced over time.

Effectiveness of Spaced Review. The finding that spaced review, or distributed practice, is more effective than missed practice for retention of meaningfully-related material is reported in several studies (Austin, 1921; Foster, 1928; Welborn, English & Killian, 1934). Reynolds and Glaser (1964) found that spaced review was more effective than an equal amount of massed practice. They compared varying amounts of massed practice with corresponding amounts of review on biology material. No differences in retention were found for amounts of repetition. However, for both immediate and delayed (three week) retention, spaced review was found to be more effective than repetition on measures of recognition, aided recall and unaided recall.

Other studies (Spitzer, 1939; Tiedman, 1948; Pathael, Fitzpatrick, & Bischof, 1953; Ausubel & Youssef, 1965; Gobson, 1965; Merrill, 1965) have also demonstrated the effectiveness of spaced reviews for meaningful materials irrespective of massed-versus-distributed practice considerations. Spitzer found that more forgetting occurred in 1 day with no review than in 63 days with 2 spaced reviews. Ausubel and Youssef found that spaced review, as compared to no review, significantly enhanced delayed retention of a reading passage. They attributed the effect of spaced review to the opportunity provided by another trial to (a) acquire new meanings and consolidate established meanings and (b) test remembered knowledge against actual content. Gagné and Bassler (1963) found an irregular pattern of recall for subordinate concepts in a learning hierarchy, and interpreted their results as implying the importance of review in any sequence of instruction.

Temporal Position of Spaced Review. If spaced reviews have the effect of improving retention, an additional problem arises as to the optimal spacing, or temporal positioning, of such reviews. The question has been investigated for meaningful verbal material by Peterson, Ellis, Toohill and Kloess (1935), Sones & Stroud (1940), and Ausubel (1966). Peterson et al., first demonstrated that for retention intervals of two, three, six, or eight weeks, retention was greater for subjects with one review (being defined as a recall-testing situation) as opposed to subjects with no review, and three times as great for subjects with two reviews as opposed to one review. In a subsequent experiment they found that a review 2 to 3 days after original learning was equally effective

for retention as a review 1 week after original learning; learning occurred at the same time for both groups and retention was measured 10 and 21 days after original learning. The experiment was repeated for reviews of one day and nine days; they were again found to be equally effective.

Sones and Stroud compared 2 types of dual reviews (testing vs. relearning) at 3 intervals--1 and 3 days, 8 and 15 days, and 15 and 17 days. In a testing review the subject simply takes a test on the learned material; in a relearning review the subject goes through the same procedures as on the day of original learning. All learning occurred at the same time and retention was measured 42 days after original learning. An interaction was found: testing reviews were more beneficial when they occurred early. Although there was a tendency for relearning reviews to be more effective when occurring later, mean differences were small (16.00, 16.41, and 16.92).

Ausubel compared a relearning review of one day with a relearning review of seven days and found a slight but non-significant difference when final retention was measured six days after the review. Ausubel used a different design in that he varied position of the learning times so that retention was measured the same amount of time after the review for both groups.

II. DESCRIPTION OF THE PRESENT STUDY

Purpose and Rationale of the Study

Unfortunately, for purposes of generalization, all of the studies in the literature concerned with review and meaningful material have utilized reading passages as the material to be retained. There is no question that acquisition of verbal information is necessary in school learning. However, it represents only one subset of educational objectives. While the effects of reviews have been inadequately investigated for verbal information, they have been totally ignored for intellectual skills. It has not even been demonstrated that reviews are beneficial for this type of learning, although it appears reasonable that they are. The critical issue involves the optimal temporal position of such reviews. The Sones and Stroud (1940) study suggested that there is a tendency for testing reviews to be more effective when placed soon after learning and for relearning reviews to be more effective at a later time. Reviews in mathematics, for example, which generally involve rule learning, are a combination of testing and relearning; the student is asked to work some problems, if he is unable to solve one he effectually "relearns" it by reviewing the steps involved in its solution. This would seem to imply that for rule using, where one review is involved, position is irrelevant.

Also, Ausubel (1966, 1968) suggests that no significant differences were found for temporal position because early and delayed reviews each have their own advantages which counterbalance each other. Early review has a strong consolidating effect on recently learned highly available material which decreases as the material becomes less available. In other words, an early review provides another opportunity to acquire potential meanings in the material partially or completely missed on the first trial. Also, an early review provides the learner with feedback for testing the correctness of the skills acquired during learning. The review for rule application would confirm correct responses, clarify ambiguities, and correct misapplications. Finally, during an early review actual rather than potential meanings interact with the residue of those meanings acquired during original learning. This enhances consolidation because the learner does not have to grasp meanings and can concentrate on trying to remember them.

The theoretical advantages of delayed review are more easily seen. After a time has passed since original learning, more has been forgotten and the learner is more highly motivated to profit from the review. Secondly, and even more important, prior forgetting conceivably has a facilitating (immunizing) effect on meaningful learning and retention because as a result of both trying and failing to remember the learner attempts to strengthen weak areas, resolve existing confusion, and increase discriminability between previously learned ideas and related new propositions. In summary then, the principal advantage of early review would appear to be its superior consolidating and sensitizing

effects, whereas the principal advantage of delayed review probably inheres in superior relearning of forgotten material, on motivational and cognitive grounds. Therefore, when comparing a group receiving one early review with a group receiving one delayed review, no differences would be expected on a retention measure due to the counterbalancing effects involved.

The present study involved two experiments, each designed to answer different questions pertinent to reviews and rule learning in mathematics. The purposes of experiment I were:

1. To determine whether a review is effective for rule application retention; and

2. To determine whether temporal position of one review is an important factor.

The purposes of experiment II were:

1. To determine the effectiveness of two reviews irrespective of temporal position, and

2. To determine the importance of temporal positioning of two reviews.

Past research suggested the following hypotheses:

1. One review will significantly enhance retention.

2. Temporal position of one review will not be a significant effect.

3. And, if indeed an early and a delayed review are equally effective due to counterbalancing of relative advantages, two reviews, one early and one delayed, will be more effective than either two reviews early after learning or two delayed reviews.

In addition to the major hypotheses, several additional questions were raised. Because of the current lack of experimental data it was felt that the questions would give some insight into the processes involved in rule learning and retention. Due to the lack of reported data, investigation of the questions was primarily exploratory.

1. What is the nature of the retention curves? Is there a point in time, following initial decline, at which the curve levels off and beyond which further forgetting is negligible?

2. What effect do reviews have on the retention curves, i.e., how do the curves differ for different treatments?

3. Are fewer examples required at the time of review than at the time of learning?

4. Are example latencies shorter at the time of review than at the time of learning?

5. Are example latencies related to the difficulty of the rule?

6. Are correct response latencies shorter at the time of review than at the time of learning?

7. Are correct answer latencies related to difficulty of the rule?

Method and Procedure

Learning Materials. Experiments I and II utilized portions of CAI materials previously described in detail (Gay, 1969), which were adapted for the present study. Both experiments involved four principles from algebra and geometry; raising an algebraic expression to an indicated power, e.g., $(2x^2y^3)^4$ = ?; determining the measure of the third angle of a triangle, e.g., Angle A = 70, Angle B = 50, Angle C = ?;

determining the exponent of a product, e.g., $a \cdot a^2 \cdot a^3 = ?$; and finding the geometric mean, e.g., the geometric mean of 16 and 4 = ?. These specific principles and their corresponding rules were chosen because they are rules which have not been previously encountered by seventh and eighth grade students (Gay, 1969) and because their entry behavior requirements are minimal. The number of previously learned skills necessary is small; to successfully learn to apply these rules, students need only skill in the basic arithmetic operations of addition, subtraction, multiplication and division.

Apparatus. The learning materials were presented by an IBM 1500 Computer-Assisted Instruction (CAI) system. Terminals for this system consist of a cathode-ray tube (CRT), a light pen, and a keyboard. The terminals were located in an air-conditioned, sound-deadened room. The CAI system administered the learning and review tasks and recorded the students' responses including response latencies.

Selection of Subjects. Subjects for experiment I were eighth grade students enrolled in the Florida State University School (N = 53). Certain segments of the total group, who were currently enrolled in an algebra class, were excluded because they had a high probability of encountering one or more of the experimental rules in their studies. Also, a low ability group was excluded; students in this group were judged by the school to be lacking in the basic entry skills required for the learning topic used in the experiment. Subjects for experiment II were seventh grade students also enrolled in the University School (N = 67). This number (67) included the entire seventh grade population

except for a low ability group which was excluded because they also lacked basic entry skills.

Absence from the learning session, any of the review sessions, or the delayed retention testing session, resulted in the subject being excluded from analysis. Such absences resulted in unequal groups.

Experimental Design. The design for both experiments followed from that used by Peterson et al. (1935) and Sones and Stroud (1940) in that the interval between learning and retention testing was equated for all groups (see Figure 1). The design used by Ausubel (1966) which equates for all groups the interval between last review and retention testing, was not chosen for two major reasons: (1) Having some students review at a time nearer to the retention test than other students was judged to be less serious than having some students learn at a time nearer to the test than other students; and (2) The Peterson-Sones and Stroud design is more like the situation in typical classroom learning, i.e., all students learn the same material at approximately the same time and retention is measured at some future time. Also, the fact that Peterson et al. and Ausubel obtained very similar results using two different methods, Peterson et al. equating learning times, Ausubel equating review times, implies that this is not a critical variable.

Experimental Procedures. Assignment to treatment groups was done using stratified random sampling. Students in both the seventh and eighth were ability grouped based on Cooperative Battery scores and teacher recommendations. The eighth grade had three ability groups (excluding algebra students) designated as mathematics 2, 3, and 4. Each

of the 2's were randomly assigned to one of the four treatment groups and similarly each of the 3's and 4's.

Experiment I

		Day								
	Group	0	1	2	3 ...	7 ...	14 ...		21	
n = 14	I	Learn	Review					Retention	Test	
n = 14	II	Learn				Review			"	
n = 13	III	Learn					Review		"	
n = 12	IV	Learn							"	

Experiment II

		Day							
	Group	0	1	2	3 ... 6		7 ...	21	
n = 18	I	Learn	Review	Review				Retention	Test
n = 17	II	Learn	Review				Review	"	
n = 17	III	Learn				Review	Review	"	
n - 15	IV	Learn						"	

Fig. 1.--Experimental designs for Experiment I and Experiment II

Students in group 4 were of extremely low ability and were judged by the school to be lacking in the basic entry behaviors required for the task. Therefore, although these students participated in the study along with the other students, their data were not analyzed. Also, if a student was absent during the learning session, the review session or at the time of final testing, he was excluded from the experimental analyses. The above procedures resulted in the following group

memberships: Group I, n = 14; Group II, n = 14; Group III, n = 13; and Group IV, n = 12.

The seventh grade had three groupings but these involved only two ability levels. In other words, there were three classes designated as mathematics 1, 2, and 3; however, mathematics 1 and 2 were heterogeneous classes which did not differ from each other but both of which differed from 3 (which contained the low-ability students). Assignment, however, was also done based on stratified random sampling of the 1's, 2's, and 3's.

The mathematics teacher was asked to select those students in group 3 who, in his judgment, were definitely equivalent to students in the eighth grade 4 group. Those students were excluded from experimental analyses. This procedure, plus further loss due to absenteeism, resulted in the following group memberships: Group I, n = 18; Group II, n = 17; Group III, n = 17; and Group IV, n = 15.

Experiment I

All subjects in experiment I learned the rules on the same day. At the beginning of the learning session all subjects were given a pretest composed of eight problems, two for each rule to be learned (see Appendix A). During the CAI learning program, the following steps were taken for each rule:

1. The rule was presented in behavioral terms. For example: "to find the geometric mean of two positive integers, multiply the two integers and find the square root of their product."

2. An example was presented which solved a problem (which was an application of the rule) in a stepwise manner. For example:

$$(5m^3n^2)^2 = ?\ m^?n^?$$
$$= (5m^3n^2)^2$$
$$= 5^2 \cdot (m^3)^2 \cdot (n^2)^2$$
$$= 25 : m^6 \cdot n^4$$
$$= 25m^6n^4$$

3. The subject was asked to solve two problems.

4. If the subject successfully solved both problems he was informed of his correctness, shown the problems and their answers and branched to the next rule; if the subject did not successfully solve both problems, he was informed of his incorrectness, shown the problems and their answers, and presented with a new example. This procedure continued until the subject had successfully solved two successive problems for each rule.

Each subject had complete control over the time taken to learn the rules; he made the decision on each program frame as to when he was ready to view the next frame. At the end of the learning session, the subjects were all administered an immediate retention test, identical in form to the pretest, but composed of different problems (see Appendix B).

Subjects in Group I returned for a review one day following original learning. Subjects in Group II returned one week following original learning, and subjects in Group III returned two weeks from the day of original learning. Subjects in Group IV received no review. When the experimental subjects returned for their review they were administered a retention test identical in format to the former tests, differing only as to the specific problems included (see Appendix C).

During the CAI review session, the following steps were taken for each rule:

1. The subject was asked to solve two problems using the rule.
2. If solved correctly, feedback was given as described for the learning session, and the subject was asked to solve two problems using the next rule.
3. If solved incorrectly (a) feedback was given as described for the learning session, (b) another example was presented, and (c) the subject was asked to solve two new problems.
4. This procedure was continued until the subject had correctly solved two successive problems for each rule.

At the conclusion of the review session every subject was given a criterion measure, again, identical in format to former tests differing only as to specific problems (see Appendix D). On the twenty-first day from original learning a delayed retention test on the learned rules was administered to all 53 subjects. This test was also a parallel form of all former tests (see Appendix G). All criterion tests presented in the appendices were given via paper and pencil; that is, they were not part of the computer-assisted sessions.

Experiment II

The learning session for experiment II was identical to that described for experiment I. Subjects in Groups I, II, and III received two reviews; subjects in Group IV received no review. Subjects in Group I received their reviews one day and two days following original learning; subjects in Group II received their reviews one day and seven days following original learning; and subjects in Group II received their reviews six days and seven days following original learning. The

procedure for review sessions was identical to that described for experiment I. On the twenty-first day from original learning, a delayed retention test was administered to all 67 subjects. All criterion measures were identical to the ones used in experiment I with the exception that two additional measures were required for the second review (see Appendices E and F).

III. ANALYSIS OF THE DATA

The data analysis is presented in three sections. First, is a discussion of several alternate scoring procedures and their corresponding reliabilities. Next is a presentation of the retention data plots and a comparison of group differences on the delayed retention criterion measure. Finally, the results of analysis of within-program variables are presented; this includes analysis of number of examples data, correct-answers data, and latency findings.

Scoring Procedures

The delayed retention criterion test was identical for both experiments. It was composed of eight problems, two for each rule. It was originally determined that if the student correctly solved either of the problems the rule would be scored as correct. It soon became apparent, however, that correct solution of one problem was not necessarily an indication of ability to apply the rule. For example, given the following two problems,

$$2x^2 \cdot 2x^2 \cdot x^3 = _____$$

$$4x^3 \cdot 3x^2 = _____$$

it was not uncommon for a student to get one of them correct when it was readily apparent that he did not know the rule. When a student

gave responses such as

$$2x^2 \cdot 2x^2 \cdot x^3 = 4x^{12}$$

$$4x^3 \cdot 3x^2 = 12x^5$$

it was an indication that he did not know whether to multiply the exponents or to add them. By multiplying in one case and adding in the other, he insured himself credit for knowing the rule under the above-described scoring procedure.

The natural alternative appeared to be a scoring procedure whereby the student must correctly solve both problems to be scored as knowing the rule. However, this procedure also had an obvious source of unreliability. For example, given the following two problems:

> Given a triangle such that the measure of angle A = 165 and the measure of angle B = 5, then the measure of angle C = _____.
>
> Given a triangle such that the measure of angle A = 110 and the measure of angle B = 25, then the measure of angle C = _____.

it was not uncommon for a student to miss one of them while at the same time knowing how to apply the rule. When a student gave responses such as

> Given a triangle such that the measure of angle A = 165 and the measure of angle B = 5, then the measure of angle C = 10.
>
> Given a triangle such that the measure of angle A = 110 and the measure of angle B = 25, then the measure of angle C = 55.

it was clear that the student knew the rule and that his error on problem two was computational; 180 - 135 = 45, obtaining an answer of 55 is the obvious computational error if the student forgets he "borrowed" one from the eight.

The only remaining logical scoring procedure was to score the test giving one point for each problem correctly solved, yielding a

score based on total number of problems correctly solved rather than total number of rules correctly applied. To summarize the three scoring methods:

Method 1. Rule is scored as correct if either problem is solved correctly (possible score range {PSR}, 0-4).

Method 2. Rule is scored as correct only if both problems are solved correctly (PSR, 0-4).

Method 3. Score is the total number of problems correctly solved (PSR, 0-8).

The KR-20 reliability was computed for each method, separately for each experiment. Experiment I yielded reliabilities of .58, .62, and .81 respectively for the three methods. Experiment II yielded reliabilities of .57, .67, and .83 respectively. Therefore, the major analyses on the delayed retention data were done on the scores resulting from application of scoring method 3.

Retention Data

For each test given, the mean and standard deviation were computed, by group and experiment (see Table 1). These means were then plotted as shown in Figure 2.

Experiment I

As Figure 2 indicates, mean scores on the pre-learn test and post-learn test were very close for all groups. Mean scores on the pre-review retention test appears to be a direct function of the amount of time that had transpired between the learning session and the review session; i.e., one day, one week, or two weeks. At the end of the respective review sessions, however, all groups had similar mean scores.

TABLE 1.--Means and standard deviations for all tests by group and experiment

	Pre-Learn	Post-Learn	Pre-Review 1	Post-Review 1	Pre-Review 2	Post-Review 2	Delayed Retention	
Experiment I								
I. One-day Review (n = 14)								
Mean	1.86	6.36	5.78	6.36			5.71	5.78*
SD	1.56	1.34	1.42	1.45			1.98	
II. One-week Review (n = 13)								
Mean	2.21	6.78	4.78	6.36			6.00	5.87*
SD	2.23	1.25	2.01	1.78			2.32	
III. Two-week Review (n = 13)								
Mean	1.85	6.31	4.15	6.00			6.08	6.15*
SD	1.91	2.21	2.73	1.87			1.98	
IV. No Review (n = 12)								
Mean	2.00	6.25					4.25	4.24*
SD	2.00	.96					2.56	
Experiment II								
I. Two Early Reviews (n = 18)								
Mean	.61	4.89	4.00	5.17	5.06	5.94	4.72	
SD	1.20	2.03	1.85	1.92	1.95	1.47	2.58	
II. One Early and One Late Review (n = 17)								
Mean	.59	5.88	4.53	5.23	5.70	6.82	6.65	
SD	.87	.93	1.74	1.92	1.53	1.07	1.11	
III. Two Late Reviews (n = 17)								
Mean	.53	5.18	2.88	5.35	5.18	5.65	5.65	
SD	.94	1.36	1.73	1.80	2.03	2.23	1.62	
IV. No Review (n = 15)								
Mean	.80	5.47					2.67	
SD	1.21	1.46					1.88	

*Adjusted

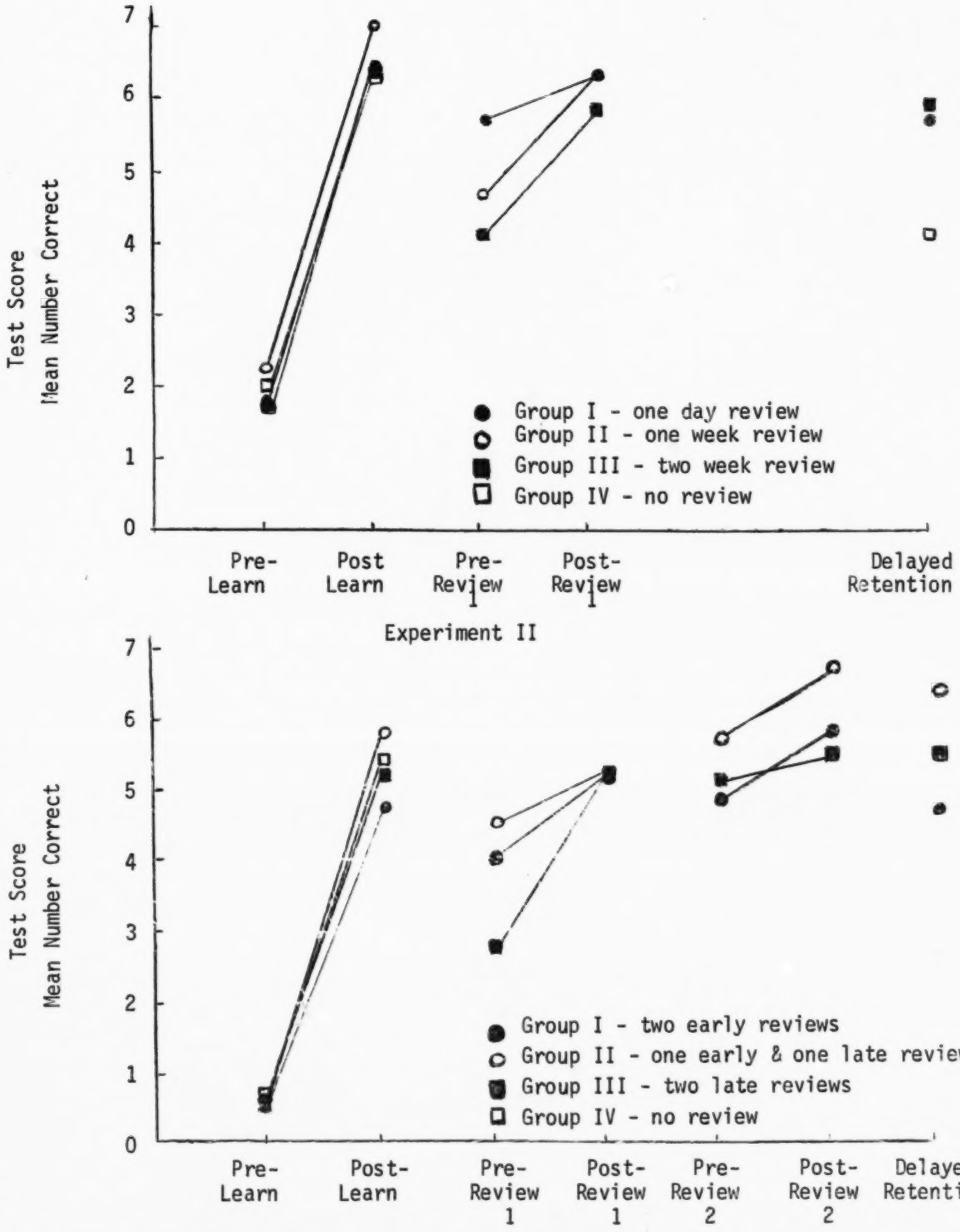

Fig. 2.--Plot of test scores for all groups before and after each learning

At the time of delayed retention testing (three weeks from the day of original learning) the mean scores were again very close. It can be seen that apparent retention losses level off after two weeks. The first retention test after the learning session is the pre-review test for the review groups and the delayed retention test for the no-review group. These means are:

Group I	One Day	5.78	(n = 14)
Group II	One Week	4.78	(n = 13)
Group III	Two Weeks	4.15	(n = 13)
Group IV	Three Weeks	4.25	(n = 12)

Although the pretest means for the groups do not appear to be too disparate, there was a correlation between pretest scores and delayed retention scores of .46 (n = 53). Therefore, a covariance analysis was performed using a General Linear Hypothesis Model with the pretest as the covariate and the delayed retention scores as the dependent variable. The specific hypotheses tested were

$$\mu 1 - \mu 2 = 0$$

$$\mu 1 + \mu 2 - 2\mu 3 = 0$$

$$\mu 1 + \mu 2 + \mu 3 - 3\mu 4 = 0$$

The adjusted treatment means were:

Group I	5.78	(n = 14)
Group II	5.87	(n = 13)
Group III	6.15	(n = 13)
Group IV	4.23	(n = 12)

While review groups retained more than the no review group ($F = 6.96$; df 1, 48; $p < .05$) the effect of temporal position of the reviews was not significant.

Experiment II

As figure 2 indicates, the pattern of retention for experiment II is nearly identical to that for experiment I up to, and including, the post-review 1 test. Group III, which did not have its first review until six days after original learning, had the lowest pre-review 1 retention score. As with experiment I, at the end of the first review session all groups had mean scores which were very close. It is interesting to note that although Groups I and III had extremely close post-review 2 scores, Group III maintained its level while Group I dropped from a mean score of 5.94 to a mean score of 4.72.

In contrast to experiment I, the pretest was not found to be significantly correlated with the delayed retention measure ($r = .16$, $n = 67$). Therefore, a simple analysis of variance was performed, using a General Linear Hypothesis Model, with the delayed retention scores as the dependent variable (see Table 2).

TABLE 2.--Analysis of variance of delayed retention scores for review and no-review subjects for Experiment II

Source	df	MS	F	P
$H0: \mu_1 - \mu_2 = 0$	1	32.39	8.83	< .01
$H0: \mu_1 + \mu_2 - 2\mu_3 = 0$	1	.53	.14	N.S.
$H0: \mu_1 + \mu_2 + \mu_3 - 3\mu_4 = 0$	1	101.07	27.56	< .01
Error	63	3.67		

The Scheffé confidence interval approach was selected for the a posteriori comparisons. It is the most conservative approach, and, according to Winer (1962), has the optimum property that the type I error is at most for any of the possible comparisons. As Table 3 shows, computation of the confidence intervals revealed that all review groups retained

TABLE 3.--Summary of Scheffé confidence interval procedure for a posteriori analysis of group means for experiment II

Groups		Scheffé Confidence Intervals
I Vs. IV	C	[.16 < μ2 - μ4 < 3.73] = .95
II Vs. IV	C	[2.06 < μ2 - μ4 < 5.90] = .95
III Vs. IV	C	[1.06 < μ3 - μ4 < 4.90] = .95

significantly more than the no review group. More interesting is the fact that Group II which received both an early and a delayed review, retained significantly more than Group I, which received two early reviews. While Group III, which received two delayed reviews, achieved a higher mean score than Group I (5.65 vs. 4.72) and Group II achieved a higher mean score than Group III (6.65 vs. 5.65), neither difference was significant, using the Scheffé technique.

Item Difficulty

One of the basic assumptions made in all of the previous analyses was that on any written test the two problems for a given rule were of equal difficulty. In an attempt to determine the soundness of this assumption, the percent of students successfully solving each problem

was computed for each group, session, and experiment. These data are presented in Table 4. As Table 4 shows, only in rare instances were the percents widely disparate, and in most cases they were either very close or identical.

Rule Mastery

Mastery generally refers to achievement of some predetermined performance criterion, e.g. 80% correct. According to this terminology, all subjects in the study achieved mastery on the four rules during learning and review sessions; each subject had to correctly solve two consecutive problems for a rule before progressing to the next, resulting in 100% performance at each session. The delayed retention test, however, indicated inability on the part of a number of students to correctly solve two problems for these same rules. In an attempt to gain some insight into the instability of exhibited mastery, an analysis was performed, by rule, which divided subjects into three categories:

1. Subjects exhibiting mastery on the delayed retention test;
2. Subjects not exhibiting mastery on the delayed retention test who did exhibit mastery on one or more of the previous written tests; and
3. Subjects not exhibiting mastery on the delayed retention test who never exhibited mastery on a written test.

These figures, expressed as percents, are presented in Table 5. With the exception of Rule 1, which was very difficult for both eighth graders (experiment I) and seventh graders (experiment II), substantial percentages of the students did achieve mastery on the delayed retention test. However, it is of importance to note that, with few exceptions, those subjects

TABLE 4.--Percent of students successfully solving each problem by group, session and experiment

Session	Pre-Learn				Post-Learn				Pre-Review 1			Post-Review 1			Pre-Review 2			Post-Review 2			Delayed Retention			
Group	I	II	III	IV	I	II	III	IV	I	II	III	I	II	III	I	II	III	I	II	III	I	II	III	IV
Problem											Experiment I													
1	00	07	00	00	57	57	57	25	50	14	31	64	71	77							36	50	54	33
2	14	21	00	08	29	36	46	42	43	29	23	64	71	69							57	50	46	25
3	50	50	46	42	86	93	77	100	71	86	77	93	93	85							79	86	97	67
4	36	43	38	50	79	86	85	83	64	93	62	86	93	77							86	93	97	67
5	36	57	46	42	93	100	85	92	93	79	69	100	86	77							100	86	97	75
6	29	29	31	17	93	100	85	92	71	71	54	71	79	69							86	86	85	66
7	07	14	23	42	93	100	100	100	100	64	54	93	79	85							64	79	77	50
8	00	00	00	00	93	93	100	83	86	43	46	71	64	62							64	71	69	42
n=	14	14	13	12	14	14	13	12	14	14	13	14	14	13							14	13	13	12
											Experiment II													
1	00	00	00	00	28	18	29	33	17	12	00	28	18	18	28	35	12	33	47	47	39	29	18	20
2	00	00	00	07	11	18	06	20	06	12	06	28	18	24	28	24	12	39	53	47	44	41	18	27
3	17	06	29	27	78	76	71	93	83	76	76	100	88	94	78	88	82	89	88	76	72	82	88	60
4	11	06	18	27	72	82	76	87	83	76	76	78	88	94	72	94	76	83	100	88	78	82	82	66
5	00	12	00	00	72	100	71	80	44	59	18	56	65	71	67	59	76	89	82	82	61	59	94	33
6	00	06	00	00	50	88	71	60	39	53	12	61	71	65	61	71	65	72	82	71	50	65	82	27
7	00	24	06	13	83	94	94	87	66	88	53	89	76	94	83	88	94	94	94	76	61	71	94	20
8	00	00	00	07	83	94	94	87	61	76	47	78	71	76	89	88	94	94	94	94	61	82	88	20
n=	18	17	17	15	18	17	17	15	18	17	17	18	17	17	18	17	17	18	17	17	18	17	17	15

TABLE 5.--Percent* of subjects achieving varying degrees of mastery** on each of the four rules for experiment I and experiment II

Ss Exhibiting Mastery on the DR*** Test				Ss Exhibiting Mastery on One or More Written Tests But Not on the DR Test				Ss Never Exhibiting Mastery			
R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4
Experiment I (n = 41)											
41.5	82.9	87.8	70.7	39.0	14.6	9.8	29.3	19.5	2.4	2.4	0.0
Experiment II (n = 52)											
25.0	82.7	63.5	73.1	21.2	13.5	28.8	23.1	53.8	3.8	7.7	3.8

* Percents do not include control subjects.
** Mastery is defined as correct solution on both problems for a given rule.
*** Delayed Retention.

who did not achieve mastery (defined as previously indicated) on the delayed retention test had achieved mastery previously on one or more written tests. The relatively high percentages of subjects in category two suggested that the current definition of mastery may be somewhat inadequate; mastery is certainly a misnomer for a performance which can not be duplicated after a brief period of time. Therefore, an analysis was performed to determine the usefulness of an alternative definition: A student achieves mastery when criterion performance is maintained over time.

Table 6 indicates the percentage of subjects achieving mastery at each session. For this analysis mastery was operationally said to be achieved at the session at which both problems were correctly solved and from which time both problems were correctly solved on all written tests, including the delayed retention test. Table 6 indicates that while a large number of subjects achieved mastery during original learning, many subjects required one, or even two, reviews to achieve it. It is also interesting to note that some subjects exhibited mastery on a pre-review measure when they did not exhibit it on the previous post-test. Whether this is a function of score unreliability or a manifestation of reminiscence is not discernible.

Group By Rule Repeated Measures Analysis. The data in Table 6 clearly show that the rules were not of equal difficulty. This suggested that a repeated measures analysis might be appropriate. In order to facilitate the analysis, subjects were randomly discarded to obtain equal n's in all groups. For experiment I, the analyses were done for groups with n = 12. For experiment II, the analyses were done for groups with n = 15. These data are presented in Table 7. The data in Table 7 agree with the previously cited data obtained from the General Linear Hypothesis Model. In addition, it can be seen that there was no interaction in either experiment, between groups and rules, although the rules effect was significant for experiment I but not for experiment II. Data presented in later tables, however, based on all the subjects in the experiment, suggest that the rule effect was the same for both experiments, and that the above finding is a function of the restricted n's used. Another possibility is

TABLE 6.--Percent* of subjects achieving mastery at each session and the cumulative percent over successive session of retention measurement, for the four rules (R1, R2, R3, R4), for experiment I and experiment II

Session	Post Learn				Pre-Review 1				Post-Review 1				Pre-Review 2				Post-Review 2				Delayed Retention			
	R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4	R1	R2	R3	R4
									Experiment I n = 41															
Percent Achieving Mastery	15	56	58	49	00	07	00	02	24	12	15	15									02	07	15	05
Cumulative Percentage	15	56	58	49	15	63	58	51	39	76	73	66									41	83	88	71
									Experiment II n = 52															
Percent Achieving Mastery	02	42	13	50	02	15	06	04	04	11	19	07	06	02	08	06	07	10	10	02	04	02	08	04
Cumulative Percentage	02	42	13	50	04	58	19	54	08	69	38	61	14	71	46	67	21	81	56	69	25	83	64	73

*To the nearest percent.

TABLE 7.--A group by rule repeated measures analysis for experiment I and experiment II

Source	df	MS	F	P
		Experiment I		
Groups	3	1.60	1.58	N.S.
Error (b)	44	1.01		
Rules	3	4.98	8.30	<.01
G x R	9	.39	.65	N.S.
Error (w)	132	.60		
		Experiment II		
Groups	3	8.06	13.66	<.01
Error (b)	56	.59		
Rules	3	1.61	2.06	N.S.
G x R	9	.15	.19	N.S.
Error (w)	168	.78		

that the results reflect a failure to meet all assumptions for a repeated measures analysis. For example, one of the major assumptions is that all measures are equally correlated. While all the rules measure mathematical ability in a general sense, two of the rules are algebraic and two are from the field of geometry. Therefore, the rules may not be equally correlated; this would constitute a direct violation of one of the primary assumptions and might account, in part, for the above findings.

Within Program Variables

Number of Examples. One of the exploratory questions raised concerned the number of examples required at the time of review as compared to the number required during the learning session; Table 8 presents this data

TABLE 8.--A comparison of the number of examples viewed at each session for experiment I and experiment II

Groups	Learn	Review 1	Review 2
Experiment I			
I--One-Day Review (n = 14)	7.8	7.6	
II--One-Week Review (n = 14)	7.6	5.9	
III--Two-Week Review (n = 13)	7.9	6.6	
Experiment II			
I--Two Early Reviews (n = 18)	10.4	10.4	4.9
II--One Early-One Late (n = 17)	8.6	9.1	6.0
III--Two Late Reviews (n = 17)	8.8	9.9	4.3

for both experiments. It can be seen that subjects required approximately the same number of examples for the first review as for the learning session. During the second review, however, the groups viewed between 35 and 50% fewer examples than at the time of learning and at the time of their first review. Looking back at Table 1, it can be noted that not only did the groups use the same number of examples, but also test scores for all groups, in both experiments, were also approximately equal at the end of the learning session and at the end of the first review session. Scores on the post-review 2 test, however, were higher than scores on the post-review 1 test. In other words, the groups used considerably fewer examples and yet improved their scores at the time of review 2.

Length of Example Latency. Two further questions were raised concerning examples:

> Are example latencies shorter at the time of review than at the time of learning? Are example latencies related to difficulty of the rule?

The answer to both question is yes (see Table 9). Before discussing these data, a word needs to be said concerning the latency data. By example, latency is meant the time elapsing between the presentation of the example and the subject's going to the next problem. It is well known that latency data, and time data in general, are positively skewed. Winer (1962) suggests that a logarithmic transformation is particularly effective in normalizing such data. In addition, as a number of subjects had a score of zero for some rules (i.e., they required no examples during a review session), all scores were transformed using the transformation

$$X' = \log (X + 1)$$

as suggested by Winer (1962). Table 9 presents mean latencies for both transformed and untransformed data. In Table 9, the first column contains indices of the difficulty of the rule based on the number of subjects correctly applying it at least once on the delayed retention measure. The other values represent mean latency for each rule by session. This data is presented by group for both experiments. In general, the order of difficulty of the rules was Rule 1, Rule 4, Rule 3, and Rule 2, with Rule 2 being the least difficult; thus the data are presented in that order.

The pattern of results for the transformed data is identical to the pattern for the untransformed data. For all six groups represented it can be seen that with few exceptions the mean latency for any given

TABLE 9.--A comparison of mean example latency and rule difficulty for experiment I and experiment II

	*DR-X Proportion Correct	$\bar{X}$ Latency Transformed			$\bar{X}$ Latency Untransformed		
		Learn	Review 1	Review 2	Learn	Review 1	Review 2
Experiment I							
Group I--One-Day Review (n = 14)							
Rule 1	.57	4.01	2.31		60.45**	15.79	
Rule 4	.64	2.47	1.30		13.25	5.30	
Rule 3	1.00	2.52	1.15		12.30	5.15	
Rule 2	.85	2.58	.90		13.75	3.72	
Group II--One-Week Review (n = 14)							
Rule 1	.50	3.86	2.43		53.69	20.64	
Rule 4	.78	2.60	1.35		14.42	7.83	
Rule 3	.86	2.35	1.01		13.19	4.15	
Rule 2	.93	2.50	.48		14.52	1.68	
Group III--Two-Week Review (n = 13)							
Rule 1	.62	3.85	2.89		54.64	23.90	
Rule 4	.77	2.60	1.60		14.44	7.25	
Rule 3	.92	2.26	1.56		9.63	5.79	
Rule 2	.92	2.42	1.02		12.09	5.21	
Experiment II							
Group I--Two Early Reviews (n = 18)							
Rule 1	.50	3.71	2.71	2.00	55.87**	21.12	13.17
Rule 4	.61	2.56	1.03	.96	14.20	5.21	5.02
Rule 3	.67	2.56	1.84	.94	13.76	6.96	3.00
Rule 2	.83	3.06	.29	.91	25.76	1.39	.79
Group II--One Early and One Late Review (n = 17)							
Rule 1	.47	4.02	2.67	1.82	83.51	20.20	10.97
Rule 4	.94	2.47	.75	.55	13.59	2.58	1.60
Rule 3	.76	2.63	1.57	.91	15.08	6.31	3.42
Rule 2	.94	2.81	.46	.06	18.09	1.57	.10
Group III--Two Late Reviews (n = 17)							
Rule 1	.35	3.88	2.78	1.92	73.02	24.31	12.12
Rule 4	.82	2.54	1.72	.79	13.11	9.45	2.74
Rule 3	.88	2.58	1.85	.66	14.16	8.69	1.66
Rule 2	.88	2.82	.40	.36	18.58	1.82	1.22

*Delayed Retention Measure
**All Untransformed Data are Expressed in Seconds

rule on the review 1 session is at least half the mean latency at the time of learning and in many cases considerably shorter. A similar statement is true concerning the change in mean latency from review 1 to review 2. In relating this to the example data, we see that at the time of review 1, subjects used the same number of examples but only viewed them, at most, half as long; at the time of review 2, however, the subjects used 35 to 50% fewer examples than at the time of review 1 and viewed them at most half as long as at the time of review 1. In general then, at the time of review 2 subjects used between 35 and 50% fewer examples and viewed them at most only 25% as long as at time of learning.

To further clarify the relationship between recall, number of examples required to reach criterion, and example latency, relearning measures were converted to savings scores and compared to percent retention scores obtained prior to the relearning. Use of savings scores is a traditional method for assessing retention. Savings scores are generally used to express a reduction in the number of trials required to reach criterion or a reduction in the amount of time taken to reach criterion. Savings scores are expressed as percents and, therefore, permit a degree of comparison between different criteria. Savings scores for both numbers of examples and example latency were computed using the following formula:

$$\text{Percent savings} = \frac{\text{learning score - review score x 100}}{\text{learning score}}$$

In every case, including savings scores for review 2, the learning score was used as the base and was assigned a value of 100%. These data are

presented in Figure 3. Figure 3 indicates, for example, that Group I, Experiment I: recalled 72% as measured by the pre-review retention test; used almost exactly the same number of examples during the review session as during the learning (indicated by a near-zero savings score); and had a mean example latency savings score of approximately 64%. Even more than Table 9, these data illustrate the tremendous savings in time that occur from session to session.

It is interesting to note that different conclusions could be reached concerning degree of retention, depending on the criterion used. If either recall scores or example latency savings scores were the index of retention, it would be concluded that retention was between 50 and 70% at the time of Review 1. If savings scores for number of examples was the index, however, it would be concluded that retention was essentially 0%. These data support the argument that retention is to a great extent a function of the method used to assess it.

Rule Difficulty and Example Latency. Another general pattern that emerges is that example latency, or the mean viewing time per example is a direct function of the difficulty of the rule as determined by success on the delayed retention measure. Further, the phenomenon becomes more pronounced with each succeeding review (see Table 9, experiment II). Furthermore, in cases where rule difficulty does not follow the general pattern, the latency still maintains its relationship with difficulty. For examples: In experiment I, Group IV, Rule 2 was more difficult than Rule 3 and correspondingly, its mean latency is longer; in experiment II, Group 2, Rule 3 was more difficult than Rule 4, and correspondingly its mean latency is longer. The only exception to the above generalization is

Experiment I

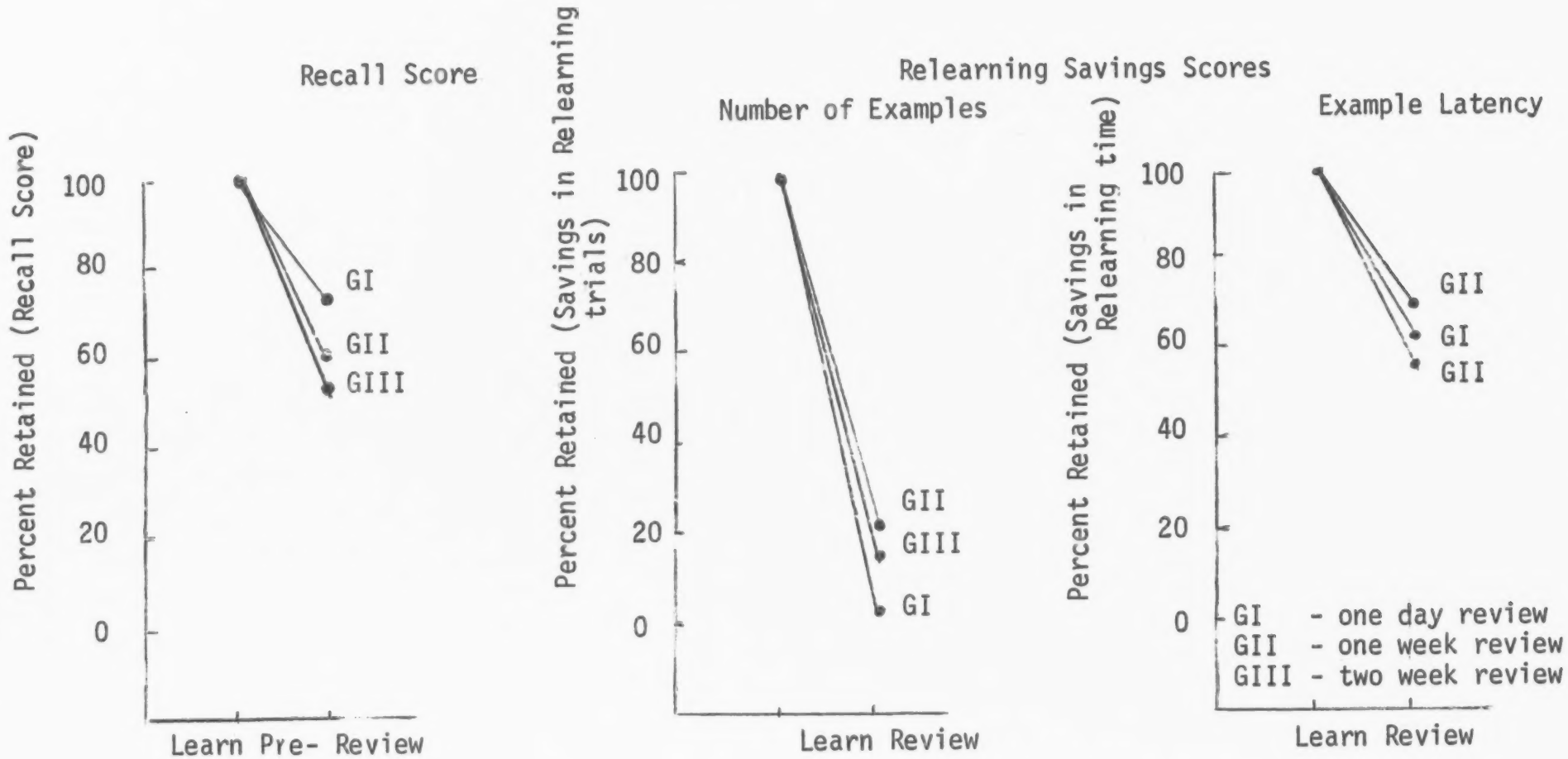

Fig. 3.--A comparison of three measures of retention for experiment I and experiment II

*Recall Score = Percent of problems correctly solved
**Savings in number of examples = (number of examples used in original learning minus the number of examples required for relearning) divided by the number of examples used in original learning.
***Savings in example latency = (mean time per example used in original learning minus the mean time per example required for relearning) divided by the mean time per example used in original learning.

Experiment II

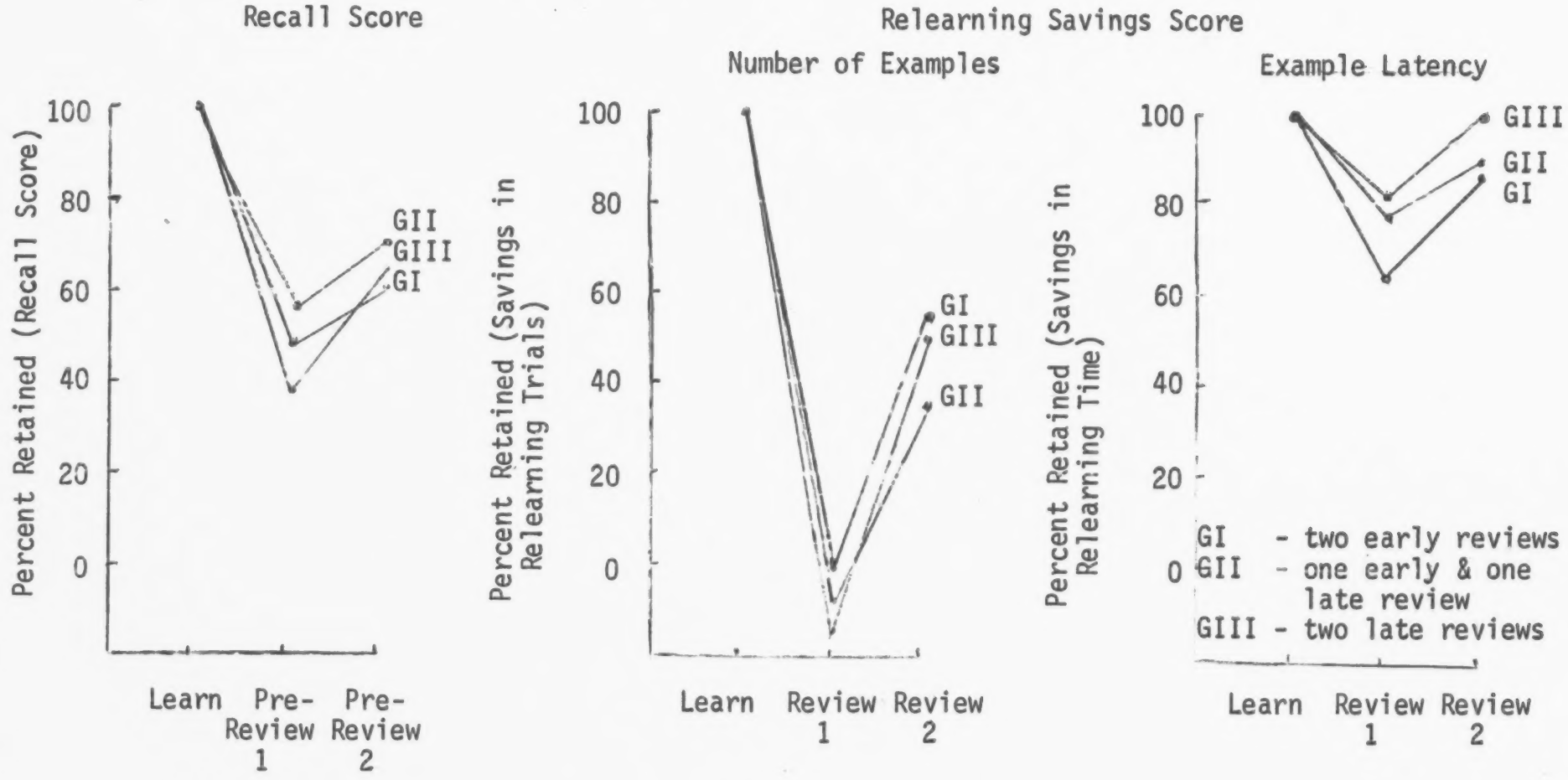

Fig. 3.-- continued

for experiment I, Group I, Rule 3; it was the easiest rule and yet its mean latency follows the same pattern as for groups where Rule 3 was the second easiest rule.

Correct Responses. The final questions raised concerned within program correct answers:

> Are correct response latencies shorter at the time of review than at the time of learning? Are correct answer latencies related to difficulty of the rule?

The answer to both questions is no (see Table 10). By correct response latency is meant the time between the presentation of the problem and the entering of the response by the subject. As Table 10 indicates correct response latencies do not get shorter at each session (nor longer). Neither is response latency related in any meaningful way to rule difficulty. There is a pattern, however; for all six groups Rule 4 had the largest mean response latencies (slight exception--experiment I, Group I). This is probably a function of the computation involved; Rule 4 is the geometric mean and requires calculation of the square root. While most of the problems involved well-known squares (e.g., 144), many students used a trial and error process to determine their roots (e.g., 9 x 9 = 81, 10 x 10 = 100, 11 x 11 = 121, 12 x 12 = 144, STOP.) It can also be seen that for all six groups Rule 3 has the shortest mean response latency. Again, this is probably a function of the small amount of computation involved. Also, in general, Rule 2 had a shorter mean response latency than Rule 1; this is consistent with the notions of computational difficulty; Rule 2 involved finding the third angle of triangle, whereas Rule 1 involved finding the exponent of the successive exponents.

TABLE 10.--A comparison of mean correct response latency and rule difficulty for experiment I and experiment II

	*DR-X Proportion Correct	Learn	Review 1	Learn 2	Learn	Review 1	Review 2
Experiment I							
Group I--One-Day Review (n = 14)							
Rule 1	.57	3.06	2.72		22.05**	15.33	
Rule 4	.64	2.98	2.59		27.46	15.28	
Rule 3	1.00	2.89	2.25		22.12	9.78	
Rule 2	.85	2.89	2.49		18.70	13.56	
Group II--One-Week Review (n = 14)							
Rule 1	.50	3.15	2.98		24.03	23.28	
Rule 4	.78	2.82	3.07		21.28	30.21	
Rule 3	.86	2.87	2.52		18.45	14.48	
Rule 2	.93	2.94	2.75		20.84	19.44	
Group III--Two-Week Review (n = 13)							
Rule 1	.62	3.15	3.02		26.11	21.69	
Rule 4	.77	2.46	3.14		13.55	28.80	
Rule 3	.92	2.76	2.53		15.46	11.91	
Rule 2	.92	2.84	2.74		18.26	15.65	
Experiment II							
Group I--Two Early Reviews (n = 18)							
Rule 1	.50	3.15	2.84	2.78	23.45**	17.61	17.67
Rule 4	.61	2.80	2.87	2.95	20.20	32.35	25.68
Rule 3	.67	2.62	2.26	2.47	13.52	11.14	11.88
Rule 2	.83	3.15	2.79	2.85	24.99	19.97	18.43
Group II--One Early and One Late Review (n = 17)							
Rule 1	.47	3.24	2.98	3.06	29.31	21.78	22.47
Rule 4	.94	2.78	3.14	3.15	18.45	37.51	26.20
Rule 3	.76	2.74	2.41	2.70	17.99	12.76	15.24
Rule 2	.94	3.16	2.62	2.96	27.50	16.40	20.62
Group III--Two Late Reviews (n = 17)							
Rule 1	.35	3.33	3.17	3.08	30.66	27.50	23.68
Rule 4	.82	2.94	3.34	3.29	23.13	36.99	31.36
Rule 3	.88	2.72	2.54	2.60	18.58	12.70	14.19
Rule 2	.88	3.00	2.74	2.84	22.59	17.18	17.80

*Delayed Retention Measure

**All Untransformed Data are Expressed in Seconds

An interesting cross-group comparison for experiment I at the time of review reveals that for all four rules, Group II had the largest mean response latencies followed by Group III, with Group I having the shortest mean response latencies. It is understandable that Group I would have the shortest latencies in that they received their review the day following learning. However, the fact that Group II has longer latencies than Group III is not so predictable and poses an interesting paradox.

In summary then, subjects required the same number of examples at the time of review, as at the time of original learning but between 35 and 50% fewer at the time of review 2. Mean example latency is approximately 50% shorter at the time of review 1 than at the time of original learning and 50% shorter at the time of review 2 than at the time of review 1. In other words, at review 1, subjects are receiving the same number of examples half as long; at review 2, subjects are receiving half as many examples one fourth as long as originally. There is no meaningful pattern to the mean correct-response latency data other than the suggestion that rules requiring more computation involve longer latencies.

A word of caution needs to be said concerning the results. All findings are based on a highly structured instructional method which is not necessarily typical of the way in which rules are generally taught. While a high degree of control is desirable from a research point of view, it necessarily detracts from the realism of the learning situation. Therefore, in addition to the usual limitations on generalizability, all results are generalizable only to comparable instructional settings.

IV. DISCUSSION

The results of experiment I indicate that the temporal position of one review is not an important factor in retention of mathematical rules. These findings are consistent with those of Peterson et al. (1935), Sones and Stroud (1940) and Ausubel (1966) involving meaningful verbal learning. Ausubel has suggested that this is a result of each review having its own relative advantages: an early review has a superior consolidating effect on the highly available material whereas a delayed review involved superior relearning of forgotten material. The results of experiment II lend support to this argument. The group which received both an early and a late review achieved a higher mean score on the delayed retention test than the group which had two early reviews and the group which had two late reviews retained approximately 20% more than the group with two early reviews; the group with both an early and a late review retained approximately 20% more than the group with two late reviews (or approximately 40% more than the group with two early reviews). This suggests that while an early and a late review each have their own unique contribution to make to retention, the contribution of the late review is greater. This is also supported by experiment I; the group with a one-week review retained slightly more than the group with a one day review and the group with a two week review retained slightly more than the group with a one-week review (this is even more pronounced for the means adjusted for pretest differences).

Although the pairs of review were differentially effective depending on their temporal position, all review groups in both experiments retained significantly more than the group which received no review. Groups with one review retained up to 45% more than the group with no review. This is highly consistent with the findings of Peterson et al. (1935) who found that after either two or three weeks a one-review group showed a superiority of 47% over a no-review group. Also consistent with the Peterson et al. study is the findings that the second review seems to be even more effective than the first. All two-review groups retained between 75% and 150% more than their corresponding no-review group. The next question to be answered concerns the optimal placing of the early and the delayed review. It would probably be agreed that the early review should come no more than a day following original learning. But what about the delayed review? When does a review stop being "early?" Is a four-day delay sufficient? Five days? The optimum position of the delayed review may be content specific, or, not really matter; the answer requires empirical investigation.

The plot of the various test means for the groups (see Figure 2) suggests several things about the retention of mathematical rules. As of the post-review 1 retention test, the pattern of retention is highly similar for both experiment I and experiment II. In both cases, pre-review 1 retention is a direct function of the amount of time which has lapsed since the time of original learning. The first review has the effect of bringing all groups back to a common level, a level which is then maintained over time if no further review is involved. As experiment I shows, retention losses appear to level off after a period between one and two

weeks following original learning; this is indicated by the fact that the group which had their review two weeks after learning had a pre-review mean retention very similar to the mean retention of the no-review group which was not tested until three weeks after learning.

The second review does not have the same effect as the first review; if it did all groups would have a similar level of delayed retention, probably somewhat higher than it would have been had they had but one review. (Note: had they had but one review, experiment 1 suggests that their final retention level would be slightly higher than the observed final retention of Group I, the group with two early reviews). In contrast, rather than bringing the groups together the second review spreads them apart, magnifying any small exisiting differences, indicating the apparent importance of the temporal positioning of those reviews.

The data on rule mastery indicate that perhaps mastery should be defined not only in terms of performance but also in terms of the permanence of that performance. A number of subjects who exhibited mastery on one or more written tests, failed to do so on the delayed retention test. The data for those subjects who did achieve and maintain mastery suggest implications for the concept of "degree of original learning," which is generally accepted to be the most important variable effecting retention. In all ways controllable, all subjects in this study achieved the same degree of original learning, varying both on the number of examples and on the time required to reach the criterion of 100% performance on each rule. And yet, as discussed previously, for each rule varying percentages of students achieved permanent mastery at this point;

each review session resulted in additional students achieving permanent mastery. This suggests that retention is a function not only of the degree of original learning, but also of the degree of relearning, or the amount and nature of review involved.

Examination of within-program variables indicates considerable time savings on each review and up to 50% savings on the number of examples required at the time of the second review. In both experiments, the same number of examples were required at the time of the first review but mean latencies per example were generally at least 50% shorter. At the same time, one review resulted in up to 45% greater retention. A second review was even more efficient; not only did the second review require up to 50% fewer examples, but the examples had mean latency saving scores of up to 95%. Further, while one review improved retention up to 45%, two reviews improved retention up to 150%. Considering the large retention losses which are generally found over time for mathematical material and given the highly hierarchical nature of mathematical material, it seems reasonable to conclude that reviews are worth the additional time, two reviews even more so than one review. That is, mathematics is a highly structured discipline and acquisition of a new concept or rule is very much dependent on availability of previously acquired relevant skills (Gagné & Paradise, 1961). Improving retention of learned material at each stage would have a cumulative positive effect on further mathematics learning.

Finally, the fact that item difficulty appears to be highly related to example latency but not to correct response latency suggests that processing time is the variable which should be studied rather

than response times. This contention is supported by the fact that response latencies appear to be confounded by a computation factor; rules which require more computation (which are not necessarily more difficult) naturally require more response time. Therefore, any attempt to relate response latencies (in a task which requires composed responses) to other variables is doomed to failure unless time to compose responses is held constant. Example latencies are apparently freer from confounding; no activity other than studying the example is reflected.

Of what import is the fact that example latencies are highly related to item difficulty? First, it suggests that the rule whose examples are studied the longest is going to be misapplied the most frequently. This implies one of two things:

1. The additional time spent viewing each example is not worth it in terms of retention. If this be the case, the optimum instructional stragegy would seem to be to tell the student with long latencies (a) not to spend so much time viewing each example, (b) to study more examples, each one for a shorter time, or (c) both.

2. Although more time is spent on the example, it is still not sufficient. The optimum instructional strategy in this case would be to tell students that the rule is very difficult and that they should study examples very carefully. An alternative approach would be to force students to view the examples longer by removing student control on presentation time.

Which of these possible explanations is correct is a question which can only be answered empirically through future research.

A further implication of the relationship between item difficulty and example latency has been put forth by Hansen (personal communication). Hansen has suggested that the fact that more complex items have longer example latencies might be related to the concept of knowledge hierarchies.

If this be the case, each succeeding level of a hierarchy would have progressively longer example latencies. If, through research these hypotheses should be found to be true, example latency could be used as a measure of the validity of a hierarchy.

In conclusion, this study has shown that additional reviews are not only highly effective in terms of delayed retention, but are also very efficient in terms of numbers of examples utilized as well as amount of time spent on those examples.

V. SUMMARY

The purpose of the present study was to investigate for mathematical rule learning one variable, spaced review, which verbal learning studies have indicated is important for retention. The effect of temporal position of review was investigated in two experiments: experiment I involved one review; experiment II involved two reviews. The methodology for both experiments paralleled that employed in meaningful verbal learning research.

Experiment I compared the relative effectiveness of one review placed either one day (Group I), one week (Group II), or two weeks (Group III) after original learning. Learning consisted of four rules presented via computer-assisted instruction. Subjects utilized varying numbers of examples, all learning to a common criterion of two consecutive correct problem solutions. The four rules involved principles from algebra and geometry; raising an algebraic expression to an indicated power, determining the measure of the third angle of a triangle, determining the exponent of the product of factors, and finding the geometric mean.

Fifty-three eighth grade students were pretested, presented with the learning task and randomly assigned to one of four groups; the four groups were the three groups described above plus a control group which received no review. The review procedure was very similar to the learning procedure only in that it began with the presentation of two new problems. All subjects relearned to a criterion of two successful problem solutions for each rule.

Three weeks from the day of original learning, all subjects were administered a delayed retention measure on the four rules. While all review groups retained considerably more than the no review group, temporal position was not significant. These findings are consistent with previous research on meaningful verbal learning.

Experiment II compared the relative effectiveness of two reviews placed either one and two days (Group I), one and seven days (Group II) or six and seven days (Group III) after original learning. Sixty-seven seventh grade students were pretested, presented with the learning task, and randomly assigned to one of four groups; the four groups were the three groups just described plus a control group which received no review. The learning and review procedures were identical to those described for experiment I, all subjects learning to a common criterion.

Three weeks from the day of learning all subjects were administered a delayed retention measure on the four rules, the same one used in experiment I. All review groups retained significantly more than the no review group. Also the group with both an early and a delayed review retained significantly more than the goup with two early reviews. This suggests that while an early and a late review each make a unique contribution to retention, the contribution of the latter is greater.

It was also found that in general, subjects in all groups required the same number of examples to reach criterion at the time of review 1 as at the time of original learning. At the time of review 1, however, example latency was approximately 50 percent shorter than at the time of learning. In contrast, at the time of review 2, subjects

required half as many examples to reach criterion as at the time of learning, with example latency being 75 percent shorter.

It was concluded that with respect to the retention of mathematical rules, one review is more effective than no review, regardless of temporal position. Optimal retention over a three week interval, however, is obtained with two reviews, one early and one delayed.

APPENDICES

APPENDIX A

LEARN PRETEST

APPENDIX A

LEARN PRETEST

NAME_____ CAI NUMBER_____

1. $(3x^3y^2)^3$ = _____

2. $(abc^3)^2$ = _____

3. Given a triangle such that the measure of Angle A = 140 and the measure of Angle B = 25, then the measure of Angle C = _____.

4. Given a triangle such that the measure of Angle A = 85 and the measure of Angle B = 75, then the measure of Angle C = _____.

5. $x \cdot x^3 \cdot x^4$ = _____.

6. $2a^2 \cdot 3a^3$ = _____.

7. The geometric mean of 8 and 8 is _____.

8. The geometric mean of 5 and 20 is _____.

APPENDIX B

POST-LEARN RETENTION TEST

APPENDIX B

POST-LEARN RETENTION TEST

NAME_____ CAI NUMBER_____

1. $(5c^2d^2)^2$ = _____.

2. $(x^2y^3z)^4$ = _____.

3. Given a triangle such that the measure of Angle A = 105 and the measure of Angle B = 60, then the measure of Angle C = _____.

4. Given a triangle such that the measure of Angle A = 95 and the meaning of Angle B = 55, then Angle C = _____.

5. $w \cdot w^2 \cdot w^3$ = _____.

6. $y^4 \cdot y^3 \cdot y^2$ = _____.

7. The geometric mean of 9 and 4 = _____.

8. The geometric mean of 18 and 2 = _____.

APPENDIX C

PRE-REVIEW 1 RETENTION TEST

APPENDIX C

PRE-REVIEW 1 RETENTION TEST

NAME_____ CAI NUMBER_____

1. $(pr^2st)^4$ = _____.

2. $(7g^3p^4)^2$ = _____.

3. Given a triangle such that the measure of Angle A = 125, and the measure of Angle B = 20, then the measure of Angle C = _____.

4. Given a triangle such that the measure of Angle A = 130 and the measure of Angle B = 15, then the measure of Angle C = _____.

5. $3m^3 \cdot m^2 \cdot m^2$ = _____.

6. $5r^3 \cdot 2r^6$ = _____.

7. The geometric mean of 9 and 1 = _____.

8. The geometric mean of 4 and 25 = _____.

APPENDIX D

POST-REVIEW 1 RETENTION TEST

APPENDIX D

POST-REVIEW 1 RETENTION TEST

NAME_____ CAI NUMBER_____

1. $(8x^{5}y^{4}z^{3})^{2}$ = _____.

2. $(b^{2}c^{6}d^{3})^{5}$ = _____.

3. Given a triangle such that the measure of Angle A = 165 and the measure of Angle B = 5, then the measure of Angle C = _____.

4. Given a triangle such that the measure of Angle A = 110 and the measure of Angle B = 25, then the measure of Angle C = _____.

5. $2x^{2} \cdot 2x^{2} \cdot x^{3}$ = _____.

6. $4x^{3} \cdot 3x^{2}$ = _____.

7. The geometric mean of 8 and 2 = _____.

8. The geometric mean of 6 and 24 = _____.

APPENDIX E

PRE-REVIEW 2 RETENTION TEST

APPENDIX E

PRE-REVIEW 2 RETENTION TEST

NAME_____ CAI NUMBER_____

1. $(2r^2s^2t^2)^3$ = _____.

2. $(k^2mn^3)^6$ = _____.

3. Given a triangle such that the measure of Angle A = 105 and the measure of Angle B = 50, then the measure of Angle C = _____.

4. Given a triangle such that the measure of Angle A = 90 and the measure of Angle B = 45, then the measure of Angle C = _____.

5. $x^6 \cdot x^7$ = _____.

6. $5b^3 \cdot 3b^4$ = _____.

7. The geometric mean of 7 and 7 = _____.

8. The geometric mean of 2 and 32 = _____.

APPENDIX F

POST-REVIEW 2 RETENTION TEST

APPENDIX F

POST-REVIEW 2 RETENTION TEST

NAME_____ CAI NUMBER_____

1. $(2f^7g^2h^4)^3$ = _____.

2. $(b^2c^6d^3)^5$ = _____.

3. Given a triangle such that the measure of Angle A = 135 and the measure of Angle B = 20, then the measure of Angle C = _____.

4. Given a triangle such that the measure of Angle A = 145 and the measure of Angle B = 30, then the measure of Angle C = _____.

5. $p^5 \cdot p^4 \cdot p^6$ = _____.

6. $2z^2 \cdot 4z^2 \cdot z^2$ = _____.

7. The geometric mean of 12 and 3 = _____.

8. The geometric mean of 3 and 48 = _____.

APPENDIX G

DELAYED RETENTION TEST

APPENDIX G

DELAYED RETENTION TEST

NAME_____ CAI NUMBER _____

1. $(6m^4n^3p^2)^2$ = _____.

2. $(xy^2z^3)^4$ = _____.

3. Given a triangle such that the measure of Angle A = 115 and the measure of Angle B = 35, then the measure of Angle C = _____.

4. Given a triangle such that the measure of Angle A = 120 and the measure of Angle B = 15, then the measure of Angle C = _____.

5. $v^7 \cdot v^5 \cdot v^3$ = _____.

6. $7x^4 \cdot 4x^6$ = _____.

7. The geometric mean of 50 and 2 = _____.

8. The geometric mean of 16 and 4 = _____.

REFERENCES

REFERENCES

Austin, S. D. M. A study of logical memory. The American Journal of Psychology, 1921, 32, 370-403.

Ausubel, D. P. Early versus delayed review in meaningful learning. Psychology In The Schools, 1966, 3, 195-198.

Ausubel, D. P. Educational psychology: A cognitive view. New York: Holt, Rinehart and Winston, Inc., 1968.

Ausubel, D. P., & Youssef, M. The effect of spaced repetition on meaningful retention. The Journal of General Psychology, 1965, 73, 147-150.

Ausubel, D. P., Robbins, L. C., & Blake, E., Jr. Retroactive inhibition and facilitation in the learning of school materials. The Journal of Educational Psychology, 1957, 48, 334-343.

Davis, R. A., & Rood, E. J. Remembering and forgetting arithmetical abilities. The Journal of Educational Psychology, 1947, 38, 216-222.

Dowling, R. M., & Braun, H. W. Retention and meaningfulness of material. The Journal of Experimental Psychology, 1957, 54(3), 213-217.

Foster, J. C. Verbal memory in preschool children. The Journal of Genetic Psychology, 1928, 35, 26-44.

Freud, H., & Cheronis, N. D. Retention in the physical science survey course. Journal of Chemical Education, 1940, 17, 289-293.

Gagné, R. M., & Bassler, O. C. Study of retention of some topics of elementary nonmetric geometry. Journal of Educational Psychology, 1963, 54(3), 123-131.

Gagné, R. M., & Paradise, N. E. Abilities and learning sets in knowledge acquisition. Psychological Monographs, 1961, 75(14), 1-23.

Gay, L. R. An investigation into the differential effectiveness for males and females, of three CAI treatments on delayed retention of mathematical concepts. Technical Report No. 12, Computer-Assisted Instruction Center, Florida State University, Tallahassee, 1969.

Gibson, J. T. The effects on retention of programmed classroom reviews. The Journal of Educational Research, 1965, 58(10), 449-452.

Kintz, B. L. Short-term retention and long-term retention as a function of practice. The Journal of Psychology, 1965, 59, 309-314.

Korn, J. H., & Lindley, R. H. Immediate memory for consonants as a function of frequency of occurrence and frequency of appearance. The Journal of Experimental Psychology, 1963, 66, 149-154.

Lahey, Sister M. F. Permanence of retention of first year algebra. The Journal of Educational Psychology, 1941, 32(6), 401-413.

Layton, E. T. The persistence of learning in elementary algebra. The Journal of Educational Psychology, 1932, 23, 46-55.

Melton, A. W. Implications of short-term memory for a general theory of memory. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 1-21.

Merrill, M. D. Correction and review on successive parts in learning a hierarchical task. Journal of Educational Psychology, 1965, 56, 225-234.

Mouly, G. J. Psychology for effective teaching. New York: Holt, Rinehart, and Winston, Inc., 1968.

Murphy, D. B. Recognition and recall as a function of frequency, shock, and individual difference variables. Doctoral dissertation, State University of Iowa. Dissertation Abstracts, 1953, 13(5), 1277-1278.

Orr, W. C. Retention as a variable in comparing programmed and conventional instructional methods. The Journal of Educational Research, 1968, 62(1), 11-13.

Oseas, L., & Underwood, B. J. Studies of distributed practice: V. Learning and retention of concepts. The Journal of Experimental Psychology, 1952, 43, 143-148

Pathael, L. L., Fitzpatrick, E. D., & Bischof, L. J. A comparison of the extent of retention accompanying the use of three teaching testing procedures. The Journal of Educational Research, 1953, 47, 65-70.

Peterson, H. A., Ellis, M., Toohill, N., & Kloess, P. Some measurements of the effects of reviews. Journal of Educational Psychology, 1935, 26, 65-72.

Peterson, M. J. Verbal response strength as a function of cultural frequency, schedule of reinforcement, and number of trials. The Journal of Experimental Psychology, 1956, 52, 371-376.

Pratt, C. C. Repetition, motivation, and recall. British Journal of Psychology, 1936, 425-429.

Pressey, S. L., Robinson, F. P., & Horrocks, J. E. Psychology in Education. New York: Harper & Row, Publishers, Inc., 1959.

Reynolds, J. H., & Glaser, R. Effects of repetition and spaced review upon retention of a complex learning task. Journal of Educational Psychology, 1964, 55(5), 297-308.

Schrepel, M., & Laslett, H. R. On the loss of knowledge by junior high pupils over the summer vacation. The Journal of Educational Psychology, 1936, 27, 299-303.

Schuell, T. J., & Keppel, G. Learning ability and retention. The Journal of Educational Psychology, 1970, 61(1), 59-65.

Shay, C. Relationship of intelligence to stepsize on a teaching machine program. The Journal of Educational Psychology, 1961, 52, 98-103.

Smeltz, J. R. Retention of learnings in high school chemistry. Science Teacher, 1956, 23, 285.

Sones, A. M., & Stroud, J. B. Review, with special reference to temporal position. Journal of Educational Psychology, 1940, 31, 665-676.

Spitzer, H. F. Studies in retention. The Journal of Educational Psychology, 1939, 30(9), 641-656.

Sterrett, M. D., & Davis, R. A. The performance of school learning: A review of studies. Educational Administration and Supervision, 1954, 40(8), 449-460.

Tiedeman, H. R. A study of classroom learning. The Journal of Educational Research, 1948, 41, 516-531.

Tulving, E. Theoretical issues in free recall. In T. R. Dixon & D. L. Horton (Eds.),Verbal behavior and general behavior theory. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1968, Chapter 1.

Tyler, R. W. What high school pupils forget. Educational Research Bulletin, 1930, 9, 490-492.

Underwood, B. J. Laboratory studies of verbal learning. In E. R. Hilgard (Ed.), Theories of learning and instruction: The sixty-third yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press, 1964.

Welborn, E. L., English, H. B., & Killian, C. D. Studies in substance memorization. The Journal of General Psychology, 1934, 11, 233-260.

Winer, B. J. Statistical principles in experimental design. New York: McGraw-Hill Book Company, Inc., 1962.

Worcester, D. A. The permanence of learning in high school subjects--algebra. The Journal of Educational Psychology, 1928, 19, 343-345.

ADDITIONAL REFERENCES

Ausubel, D. P., & Blake, E., Jr. Proactive inhibition in the forgetting of meaningful school learning. The Journal of Educational Research, 1958, 52 (4), 145-149.

Bowers, N. D. Meaningful learning and retention: Task and method variables. Review of Educational Research, 1961, 31, 522-534.

Briggs, L. J., & Hamilton, N. R. Meaningful learning and retention: Practice and feedback variables. Review of Educational Research, 1964, 34, 545-558.

Bunch, M. E. The measurement of retention by the relearning method. Psychological Review, 1941, 48, 450-456.

Buschke, H., & Lim, H. Short-term storage of repetitions of two items. Psychonomic Science, 1967, 7, 277-278.

Cosper, R., & Kephart, N. C. Retention of reading skills. The Journal of Educational Research, 1955, 49, 211-216.

Gagné, R. M., & Brown, L. R. Some factors in the programming of conceptual learning. The Journal of Experimental Psychology, 1961, 62 (4), 313-321.

Gagné, R. M., & Wiegand, V. K. Some factors in children's learning and retention of concrete rules. Journal of Educational Psychology, 1968, 59, 355-361.

Gillette, A. L. Learning and retention: A comparison of three experimental procedures. Archives of Psychology, 1936, 198, 1-56.

Guthrie, J. T. Expository instruction versus a discovery method. The Journal of Educational Psychology, 1967, 58 (1), 45-49.

Kittel, J. E. An experimental study of the effect of external direction during learning on transfer and retention of principles. Journal of Educational Psychology, 1957, 48, 391-405.

Klausmeier, H. J., & Feldhusen, J. F. Retention in arithmetic among children of low, average, and high intelligence at 117 months of age. The Journal of Educational Psychology, 1959, 50 (2), 88-92.

Krumboltz, J. D. Meaningful learning and retention: Practice and reinforcement variables. Review of Educational Research, 1961, 31, 535-546.

Little, E. B. Pretest and retest scores in retention calculation. The Journal of Experimental Education, 1960, 29 (2), 161-167.

Mecon, L. J. Concept learning and retention in mathematics. The Journal of Experimental Education, 1967, 36 (1), 51-57.

Schaaf, O. Student discovery of algebraic principles as a means of developing ability to generalize. The Mathematics Teacher, 1955, 48, 324-327.

Schuster, A. H., & Pigge, F. L. Retention of meaningful teaching. The Arithmetic Teacher, 1965, 12, 24-31.

Shiffrin, R. M., & Atkinson, R. C. Storage and retrieval processes in long-term memory. Psychological Review, 1969, 76 (2), 179-193.

Stroud, J. B., & Schoer, L. Individual differences in memory. The Journal of Educational Psychology, 1959, 50 (6), 285-292.

Underwood, B. J. Ten years of massed practice on distributed practice. Psychological Review, 1961, 68 (4), 229-247.

Welborn, E. L., & English, H. Logical learning and retention: A general review of experiments with meaningful verbal materials. The Psychological Bulletins, 1937, 34 (1), 1-20.

Wittrock, M. C. Verbal stimuli in concept formation: Learning by discovery. Journal of Educational Psychology, 1963, 54 (4), 183-190.

Wittrock, M. C., & Twelker, P. A. Prompting and feedback in the learning retention, and transfer of concepts. The British Journal of Educational Psychology, 1964, 34, 10-18.

Wodtke, K. H. On the assessment of retention effects in educational experiments. The Journal of Experimental Education, 1967, 35 (4), 28-36.

Word, A. H., & Davis, R. A. Acquisition and retention of factual information in seventh-grade general science during a semester of eighteen weeks. The Journal of Educational Psychology, 1939, 30, 116-125.

VITA

VITA
November, 1970

Biographical Data

Name: Lorraine R. Gay
Born: January 6, 1944, Boston, Massachusetts
Marital Status: Married, no children
Social Security No.: 265-74-9051
Address: Computer-Assisted Instruction Center
Tully Building, Room 1-A
Florida State University
Tallahassee, Florida 32306
Phone: 904-599-3660

Academic History

1962-1966 B.Ed., cum laude, (1966), Mathematics, University of Miami, Miami, Florida

1967-1968 M.Ed., (1968), Educational Research and Statistics, University of Miami, Miami, Florida

1968-1970 Ph.D., (1970), Educational Research, Florida State University, Tallahassee, Florida

Positions Held

9/70-present Research Associate, Board of Regents, Division of Academic Affairs, State of Florida University System, Tallahassee, Florida

Project Director, High Priority Education Project

6/70-9/70 Research Assistant to Robert M. Gagne', Department of Educational Research, Florida State University, Tallahassee, Florida

Design of, and development of materials for, EDR 577--"Learning Theories Relevant to Education," a graduate course offered by the Department of Educational Research

9/69-9/70 Research Assistant, Computer-Assisted Instruction Center, Florida State University, Tallahassee, Florida

A major responsibility in executing an NSF-sponsored research study utilizing the IBM 1500 instructional system "A Comparison of Four Feedback Intervals for Three Types of Mathematics Instruction: Definitions, Algorithms, and Proofs"
Responsibility for data analysis of various ongoing research projects

9/68-9/69 U.S. Office of Education Fellow in Computer-Assisted Instruction and Research Assistant, Florida State University, Tallahassee, Florida

Participated in development of CAI materials on Boolean Algebra for a Naval Training Device Center training program, Orlando, Florida

Conducted a study "An investigation into the Differential Effectiveness for Males and Females, of Three CAI Treatments on Delayed Retention of Mathematical Concepts"

6/68-9/68 Research Associate, University of Miami faculty, Miami, Florida

Held position as Research Associate on the government funded Florida Migratory Child Research Project. Directed participating principals and teachers throughout the state of Florida. Wrote the chapter on the education of migratory children in Florida for final report.

9/67-6/68 Research Assistant, University of Miami, Miami, Florida

Participated in the evaluation of modular scheduling at North Miami High School, Miami, Florida

9/66-6/67 High School mathematics teacher at Miami Central High School, Miami, Florida

Taught algebra, geometry, and basic mathematics for seniors

Professional Affiliation

American Educational Research Association

Present Research Interests

Factors effecting retention of intellectual skills, e.g., rule learning in mathematics. Investigation of factors such as optimal number of examples per rule and the effectiveness and temporal position of reviews, utilizing computer-assisted instruction

Publications

Gay, Lorraine R. An investigation into the differential effectiveness for males and females, of three CAI treatments on delayed retention of mathematical concepts. Technical Report No. 12, Computer-

Assisted Instruction Center, Florida State University November, 1969. Paper read at the American Educational Research Association Convention, Minneapolis, 1970.

Gay, Lorraine R. Temporal position of reviews and its effect on retention of mathematical rules, in preparation. Paper to be read at the American Educational Research Association convention, New York, 1971.

Hicken, James E., Gay, Lorraine, R., et al. An assessment of a senior high modular scheduling program. Florida Educational Research and Development Council, June, 1968.

Kleinhert, Jack E., Gay, Lorraine, R., et al. Final Report, Phase I, Florida Migratory Research Project, September, 1968.

Paper Presented at Professional Meeting

Gay, Lorraine R. Use of a retention index for mathematics instruction. Paper read at the American Educational Research Association convention, Minneapolis, 1970.